DA4 Assignment: 2: Panel practice – João Reis

Reis-Joao/Assignment2_Data4: This repository contains the data and the code for the second assignment of Data4 made by João Reis (github.com)

Data Collection and Treatment

In the second assignment of Data 4 we were asked to study the impact of economic activity on the level of CO_2 emissions. To do it, I have downloaded the following data for the 1992-2018:

Two main variables:

- Total CO_2 emissions (thousand metric tons of CO_2 excluding Land-Use Change and Forestry)
- GDP, PPP (constant 2017 international $)

Four possible confounders:

- Population, total
- Urban population, total
- Forest area (sq. km)
- Renewable energy consumption (% of total final energy consumption)

After downloading and merging the data into one dataset (Stata Do-file in GitHub repo), I generated the CO_2 emissions and the GDP per capita, dividing their total values by the population. After that, I checked for coverage in the main variables to balance my panel data. I realized that a considerable group of countries were missing GDP observations from 1992 to 1994, and that such group was not random (mainly ex-Soviet/Yugoslavia countries). To avoid bias in the estimation, and to keep those observations (that I considered important while studying the CO_2 – economic activity relationship, due to the specifics of the planned economy and the transition process), I decided to shorten my time span and start the analysis in 1995. In the confounders, since only 3 countries did not have perfect coverage, and given that those countries were fairly random, I dropped them.

Finally, before starting my regressions, I have generated and labelled all the possible variables necessary to do it (natural logarithms, except for renewable energy – since it is in % - and first differences), and I did a summary of those same variables to verify that there were no observable anomalies. By curiosity, I also draw a twoway tsline (a line graph where the x axis is the time variable) with both GDP per capita and CO_2 emissions per capita, to have a general idea of the expected relationship between both, and to once again verify that nothing strange was happening with the data (those graphs can be accessed in the GitHub repository, or by running the Stata code). The countries used in the graphs were arbitrary (based on my curiosity), and one can change it at their pleasure.

Regression Models

After the data management, and believing on its credibility, I proceeded to compute the regressions asked in the assignment.

1. **Cross-section OLS for one year**

To compute the OLS for one year I choose 1995. I computed all kind of models (level-level, level-log, log-level, and log-log). One can access them all, but to interpret I will focus on the log-log level model. The model comes as

$$\ln(CO2\ Emissions\ per\ Capita_i)^E = \alpha + \beta \ln\left(GDP\ per\ capita_i\right)$$

and the output is:

```
. reg ln_CO2_emissions_per_capita ln_gdp_per_capita if year == 1995

      Source |       SS           df       MS      Number of obs   =       167
-------------+----------------------------------   F(1, 165)       =     56.84
       Model |   287.55151         1   287.55151   Prob > F        =    0.0000
    Residual |  834.691023       165  5.05873347   R-squared       =    0.2562
-------------+----------------------------------   Adj R-squared   =    0.2517
       Total |  1122.24253       166  6.76049718   Root MSE        =    2.2492

------------------------------------------------------------------------------
ln_CO2_emission~a | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
ln_gdp_per_capita |   1.120261   .1485875     7.54   0.000     .8268829    1.413639
            _cons |  -.9333484   1.327044    -0.70   0.483    -3.553525    1.686828
------------------------------------------------------------------------------
```

$\alpha = -0.93 \rightarrow$ the constant is not directly interpretable in a log-log model since it tells us how much the natural logarithm of CO2 emissions per capita is expected to be when the natural logarithm of GDP per capita equals 0 (and so GDP per capita equals 1).

$\beta = 1.12 \rightarrow$ the slope coefficient in a log-log model has the interpretation of an elasticity. In this case, we can say that, in 1995, CO2 emissions per capita were expected to be 1.12% higher, on average, for observations with 1% higher GDP per capita. We can also observe that the slope coefficient is significant at any level, and the 95% confidence interval completely discards the possibility of our coefficient to be equal to zero.

2. Cross-section OLS for a year of your choice

On the second OLS regression, I chose the year 2018. Once again, I computed all the models, but I am going to focus on the log-log one. The model is the same, with the difference being in the year of the observations considered.

$$\ln(CO2\ Emissions\ per\ Capita_i)^E = \alpha + \beta \ln\left(GDP\ per\ capita_i\right)$$

The output is:

```
. reg ln_CO2_emissions ln_gdp_per_capita if year == 2018

      Source |       SS           df       MS      Number of obs   =       167
-------------+----------------------------------   F(1, 165)       =     50.56
       Model |   223.46045         1   223.46045   Prob > F        =    0.0000
    Residual |  729.222646       165  4.41953118   R-squared       =    0.2346
-------------+----------------------------------   Adj R-squared   =    0.2299
       Total |  952.683095       166  5.73905479   Root MSE        =    2.1023

------------------------------------------------------------------------------
ln_CO2_emission~a | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
ln_gdp_per_capita |   1.039222    .146149     7.11   0.000     .7506582    1.327785
            _cons |  -.1335297   1.382369    -0.10   0.923    -2.862941    2.595882
------------------------------------------------------------------------------
```

$\alpha = -0.133 \rightarrow$ as in the previous regression, such coefficient is not directly interpretable.

$\beta = 1.03 \rightarrow$ We can say that, in 2018, CO2 emissions per capita were expected to be 1.03% higher, on average, for observations with 1% higher GDP per capita. We can also observe that the slope coefficient is significant at any level, and the 95% confidence interval completely discards the possibility of our coefficient to be equal to zero.

Overall, in the beginning and in the end of our time period, there is a statistically significant relationship between CO2 emissions and GDP per capita. To deepen the causality of such relation, we are going to analyse the next models.

3. First difference model, with time trend, no lags

The first difference model with time trend and no lags comes as following:

$$\Delta \ln(CO2\ Emissions\ per\ Capita)_{it}^{E} = \alpha + \beta_0 \Delta \ln(GDP\ per\ Capita)_{it} + \beta_1 Year$$

The outcome is:

```
. reg d_ln_CO2_emissions_per_capita d_ln_gdp_per_capita c.year [w=population], cluster(countryid)
(analytic weights assumed)
(sum of wgt is 147,132,356,669)

Linear regression                               Number of obs     =      3,841
                                                F(2, 166)         =      47.12
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1334
                                                Root MSE          =     .05904

                                (Std. err. adjusted for 167 clusters in countryid)
------------------------------------------------------------------------------------
                   |               Robust
d_ln_CO2_emission~a | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------------+----------------------------------------------------------------
d_ln_gdp_per_capita |   .6107395   .0634788     9.62   0.000     .4854098    .7360693
              year |  -.0003096   .0001983    -1.56   0.120    -.0007012    .0000819
             _cons |    .633996   .3968073     1.60   0.112    -.1494435    1.417436
------------------------------------------------------------------------------------
```

Since the difference of logarithms is approximately a percentage change, we have that:

$\alpha = 0.63 \rightarrow$ it is the trend in the CO2 emissions per capita. In years or in countries where the GDP per capita has not changed, CO2 emissions per capita were expected to increase by 0.63%.

$\beta_0 = 0.61 \rightarrow$ the average change in CO2 emissions per capita is expected to be 0.61% higher in countries that, during the same period, have experienced a 1% higher change in GDP per capita. In other words, either we compare the same country in different years, or different countries in the same year, CO2 emissions per capita are expected to change 0.61% more when the GDP per capita increases by 1% more.

$\beta_1 = -0.003 \rightarrow$ this coefficient captures a linear time trend of the CO2 emissions per capita. Every consecutive year has a negative impact of 0.003 percentage points in the change of CO2 emissions per capita. Such coefficient captures the impact of the unknown confounders (or known but not measured) in the CO2 emissions per capita over time. Later I will add some confounders and we can compare this coefficient to the same one in the new regression.

That said, the coefficient of the linear time trend is not statistically significant. The slope coefficient, however, is highly significant (at any level of confidence), reinforcing that there is at least a positive correlation between the change in CO2 emissions per capita and the change in GDP per capita.

4. First difference model, with time trend, 2 years lags

The first difference model with time trend and 2 years lags come as following:

$$\Delta \ln(CO2\ Emissions\ per\ Capita)_{it}{}^{E} = \alpha + \beta_0 \Delta \ln(GDP\ per\ Capita)_{it} + \beta_1 \Delta \ln(GDP\ per\ Capita)_{i,t-1} + \beta_2 \Delta \ln(GDP\ per\ Capita)_{i,t-2} + \beta_3 Year$$

The outcome is:

```
. reg d_ln_CO2_emissions_per_capita L(0/2).d_ln_gdp_per_capita c.year [w=population], cluster(countryid)
(analytic weights assumed)
(sum of wgt is 135,940,220,195)

Linear regression                               Number of obs     =      3,507
                                                F(4, 166)         =      32.99
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1486
                                                Root MSE          =     .05886

                             (Std. err. adjusted for 167 clusters in countryid)
------------------------------------------------------------------------------
                   |               Robust
d_ln_CO2_emission~a | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------------+----------------------------------------------------------
d_ln_gdp_per_capita |
               --. |   .6451099   .1096938     5.88   0.000     .4285351    .8616847
               L1. |   .0114929   .0645167     0.18   0.859    -.1158863     .138872
               L2. |  -.0123943   .0662474    -0.19   0.852    -.1431905    .1184018
                   |
              year |  -.0004719   .0004221    -1.12   0.265    -.0013053    .0003616
             _cons |   .9589589   .8465125     1.13   0.259    -.7123597    2.630277
------------------------------------------------------------------------------
```

Both the constant and the linear time trend coefficient are still not statistically significant.

$\beta_0 = 0.645 \rightarrow$ the average change in CO2 emissions per capita is expected to be 0.645% higher in countries that, during the same period, have experienced a 1% higher change in GDP per capita. In other words, either we compare the same country in different years, or different countries in the same year, CO2 emissions per capita are expected to change 0.645% more when the GDP per capita increases by 1% more.

$\beta_1 = 0.01 \rightarrow$ the average change in CO2 emissions per capita is expected to be 0.01% higher in countries that, in the year before, have experienced a 1% higher change in GDP per capita. In other words, either we compare the same country in different years, or different countries in the same year, CO2 emissions per capita are expected to change 0.61% more when the GDP per capita increases by 1% more in the year before.

$\beta_2 = -0.01 \rightarrow$ the same interpretation as before, but with the change in GDP per capita occurring two years before the period of analyse to the change in the CO2 emissions per capita.

However, none of the lags is statistically significant, both with an extremely high P-value.

5. First difference model, with time trend, 6 years lags

The first difference model with time trend and 2 years lags come as following:

$$
\begin{aligned}
\Delta \ln(CO2\ Emissions\ per\ Capita)_{it}{}^{E} &= \alpha + \beta_0 \Delta \ln(GDP\ per\ Capita)_{it} + \beta_1 \Delta \ln(GDP\ per\ Capita)_{i,t-1} \\
&+ \beta_2 \Delta \ln(GDP\ per\ Capita)_{i,t-2} + \beta_3 \Delta \ln(GDP\ per\ Capita)_{i,t-3} \\
&+ \beta_4 \Delta \ln(GDP\ per\ Capita)_{i,t-4} + \beta_5 \Delta \ln(GDP\ per\ Capita)_{i,t-5} \\
&+ \beta_6 \Delta \ln(GDP\ per\ Capita)_{i,t-6} + \beta_7 Year
\end{aligned}
$$

The outcome is:

```
. reg d_ln_CO2_emissions_per_capita L(0/6).d_ln_gdp_per_capita c.year [w=population], cluster(countryid)
(analytic weights assumed)
(sum of wgt is 112,637,556,598)

Linear regression                               Number of obs     =      2,839
                                                F(8, 166)         =      25.94
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1860
                                                Root MSE          =      .0578

                                  (Std. err. adjusted for 167 clusters in countryid)
------------------------------------------------------------------------------------
                   |               Robust
d_ln_CO2_emission~a | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------------+----------------------------------------------------------------
d_ln_gdp_per_capita |
               --. |   .6851871   .1225803     5.59   0.000     .4431698    .9272044
               L1. |   .0063003   .0904864     0.07   0.945    -.1723522    .1849528
               L2. |  -.0438065   .0733836    -0.60   0.551    -.1886921    .1010791
               L3. |   .0874414   .0498134     1.76   0.081     -.010908    .1857908
               L4. |   .0642223   .0319594     2.01   0.046     .0011231    .1273216
               L5. |  -.0607893   .0722242    -0.84   0.401    -.2033856     .081807
               L6. |  -.0547872   .0664653    -0.82   0.411    -.1860136    .0764392
                   |
              year |  -.0013078   .0009881    -1.32   0.187    -.0032587    .0006431
             _cons |   2.639705   1.987558     1.33   0.186    -1.284445    6.563855
------------------------------------------------------------------------------------
```

The interpretation of the coefficients is the same as before, adapting to the lag we are considering. I would just like to highlight some curiosities: the coefficient of the contemporaneous observation has increased with the increase in the number of lags used, and it has always been strongly significant. So, even though lags are not statistically significant per se, they seem to strengthen the relationship between a contemporaneous change in CO2 emissions and GDP per capita.

When adding so many lags, one can ask the statistical significance of the cumulative coefficient.

```
. test d_ln_gdp_per_capita + L.d_ln_gdp_per_capita + L2.d_ln_gdp_per_capita + L3.d_ln_gdp_per_capita + L4.d_ln_gdp_per_capita + L5.d_ln_gdp_per_capita + L6.d_ln_gdp_per_capita = 0

 ( 1)  d_ln_gdp_per_capita + L.d_ln_gdp_per_capita + L2.d_ln_gdp_per_capita + L3.d_ln_gdp_per_capita + L4.d_ln_gdp_per_capita + L5.d_ln_gdp_per_capita + L6.d_ln_gdp_per_capita = 0

       F(  1,   166) =   64.41
            Prob > F =    0.0000
```

After computing an F-test (since I have not computed the regression with second differences, my cumulative effect is captured by the sum of the different lags coefficients, and so I am testing multiple coefficients simultaneously), and given that the Prob > F is less than 0.05 and 0.01, we can reject the null hypothesis at 5% and even 1% significance level. That means that our cumulative effect is strongly statistically significant.

6. Fixed effects model with time and country fixed effects

After computing the First Difference models, we were asked to compute a Fixed Effects Model with time and country fixed effects. Such model comes as:

$$ln(CO2\ Per\ Capita)_{it}{}^{E} = \alpha_i + \beta_0\ ln(GDP\ per\ capita)_{it}$$

In Stata, the outcome of a fixed effects regression comes as:

```
. xtreg ln_CO2_emissions_per_capita ln_gdp_per_capita i.year [w=average_population], fe cluster(countryid)
(analytic weights assumed)

Fixed-effects (within) regression               Number of obs     =      4,008
Group variable: countryid                       Number of groups  =        167

R-squared:                                      Obs per group:
     Within  = 0.7800                                         min =         24
     Between = 0.2651                                         avg =       24.0
     Overall = 0.2636                                         max =         24

                                                F(24,166)         =     180.72
corr(u_i, Xb) = 0.2106                          Prob > F          =     0.0000

                                 (Std. err. adjusted for 167 clusters in countryid)
------------------------------------------------------------------------------------
                  |               Robust
ln_CO2_emission~a | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
------------------+-----------------------------------------------------------------
ln_gdp_per_capita |   .6999095   .0821774     8.52   0.000     .5376619     .862157
                  |
             year |
             1996 |   .0036851    .013296     0.28   0.782     -.022566    .0299363
             1997 |    .016687   .0230351     0.72   0.470    -.0287926    .0621665
             1998 |   .0223644   .0276351     0.81   0.420    -.0321972     .076926
             1999 |   .0173378   .0414302     0.42   0.676    -.0644602    .0991358
             2000 |   .0265262   .0414487     0.64   0.523    -.0553083    .1083607
             2001 |   .0368419   .0430427     0.86   0.393    -.0481399    .1218237
             2002 |   .0467167   .0424662     1.10   0.273    -.0371267    .1305601
             2003 |   .0748583   .0364134     2.06   0.041     .0029652    .1467514
             2004 |   .1024251   .0378879     2.70   0.008     .0276208    .1772293
             2005 |    .118899   .0434147     2.74   0.007     .0331827    .2046152
             2006 |    .125644   .0506581     2.48   0.014     .0256268    .2256611
             2007 |   .1405519   .0555094     2.53   0.012     .0309564    .2501474
             2008 |   .1425931   .0591166     2.41   0.017     .0258758    .2593104
             2009 |   .1491283   .0666422     2.24   0.027     .0175528    .2807039
             2010 |     .17593   .0686048     2.56   0.011     .0404796    .3113805
             2011 |   .2002863   .0750438     2.67   0.008      .052123    .3484496
             2012 |   .2113855   .0805515     2.62   0.009      .052348     .370423
             2013 |   .2161016   .0823494     2.62   0.009     .0535145    .3786888
             2014 |    .221836   .0885769     2.50   0.013     .0469536    .3967185
             2015 |   .2052982   .0906252     2.27   0.025     .0263717    .3842248
             2016 |   .1929973   .0928465     2.08   0.039     .0096852    .3763094
             2017 |    .197806   .0983966     2.01   0.046      .003536     .392076
             2018 |   .2046468   .1014588     2.02   0.045     .0043308    .4049628
                  |
            _cons |   6.595547   .7053471     9.35   0.000      5.20294    7.988155
------------------+-----------------------------------------------------------------
          sigma_u |  2.1796757
          sigma_e |  .16883997
              rho |  .99403557   (fraction of variance due to u_i)
------------------------------------------------------------------------------------
```

First, we can observe that Stata just give us one constant in the output, even though fixed effects presume a different constant for each cross-sectional observation (country, in this case). That is due to the high number of constants that it would represent. So, the constant reported by Stata is the sum between a constant term and the average of each country fixed effect (and so not meaningful). Second, we have the coefficient for each year in the time period. Their individual coefficient is not our concern, but we can see that both significance and value have increased with time. Since those coefficients capture the effect of time in the CO2 emissions (again, that might capture the impact of unknown or unobserved confounders), we can say that CO2 emissions per capita have been increased in the last few years, controlling for GDP per capita (same conclusion as with the time trend captured in the FD model). Finally, our β_0 coefficient equals 0.6999. It means that, where and when GDP per capita is higher by 1% compared to its mean within countries, CO2 emissions per capita are expected to be 0.69% larger than their mean within those same countries. In other words, if we compare two observations, the one with 1% higher GDP per capita compared to its country specific mean has, on average, 0.69% larger CO2 emissions per capita, compared to their country specific mean.

7. Long difference model

Finally, we were asked to compute a Long Difference Model. A Long Difference Model is a model based on the difference between two observations that distance each other a reasonable long period. In this case, I used the difference between the 2018 and the 1995 observations. The model come as:

$$[ln(CO2\ Per\ Capita)_{2018} - \ln(CO2\ Per\ Capita)_{1995}]^{E} = \alpha + \beta_0\ [ln(GDP\ per\ capita)_{2018} - \ln(GDP\ per\ capita)_{1995}]$$

The output is:

```
. reg ld_ln_CO2_emissions_per_capita ld_ln_gdp_per_capita [w=average_population], cluster(countryid)
(analytic weights assumed)
(sum of wgt is 152,611,504,441.8)

Linear regression                               Number of obs     =      4,008
                                                F(1, 166)         =      21.48
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3420
                                                Root MSE          =     .47961

                                  (Std. err. adjusted for 167 clusters in countryid)
------------------------------------------------------------------------------------
                     |               Robust
ld_ln_CO2_emission~a | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
---------------------+--------------------------------------------------------------
ld_ln_gdp_per_capita |    .592311   .1277971     4.63   0.000     .3399938    .8446283
               _cons |   .3036271   .1184195     2.56   0.011     .0698247    .5374295
------------------------------------------------------------------------------------
```

$\alpha = 0.3$ → on average, the change in CO2 emissions per capita between 1995 and 2018, had the GDP per capita been constant in the same period, would be 0.3%.

$\beta_0 = 0.59$ → the average change in CO2 emissions per capita between 1995 and 2008 is expected to be 0.59% higher in countries that, during the same period, have experienced a 1% higher change in GDP per capita. In other words, either we compare the same country in different years, or different countries in the same year, CO2 emissions per capita between 1995 and 2018 are expected to have changed 0.59% more when the GDP per capita increased by 1% more in the same period.

Both coefficients are, once again, statistically significant at 5% level of significance (the slope coefficient is significant at any level).

Possible Mechanisms and Confounders

The previous models give us already a very solid background to state that the CO2 emissions and the GDP per capita are strongly correlated. However, possible confounders are not hard to imagine, and to get closer to a causal relationship, their inclusion in the model is necessary. As confounders I have chosen four variables: population, urban population, forest area and renewable energy. I have chosen those confounders since I believe that they are possible common confounders and, more important, are unit and time specific, and so not addressed directly by the regressed models. I expect that more population and urban population cause both CO2 emissions and GDP per capita to growth. Concerning forest area, I believe the opposite happens. Finally, I expect renewable energy to have a negative impact on CO2 emissions, but I have no previous knowledge on its impact sign on GDP per capita (both seem arguable, depending on country's characteristics).

After regress the previous models with confounders (Annex and GitHub repository), it is possible to observe that both population and renewable energy are significant and have the expected sign in all the regressions (the former positively correlated with CO2 emissions, the latter negatively). Concerning the forest area, and against my predictions, it is just significant in the OLS models, with a positive sign. Finally, urban population is positive, as expected, but only strongly significant in the FE model. In what we are concerned the most, the introduction of confounders highly reduced the coefficient of GDP per capita. In the OLS models, that coefficient was 1.12 in 1995 and 1.03 in 2018, dropped to 0.675 and 797 respectively. In the First Differences Model the coefficient was 0.61, dropped to 0.355. In the Fixed Effects, the drop was from 0.69 to 0.265. And in the Long Difference Model, the coefficient came from 0.59 to 0.350. Overall, the coefficients in every model dropped around half of their initial value with the introduction of these confounders, meaning that a lot of the previous stated relationship between GDP and CO2 emissions had them in the background. One curiosity is that the coefficient of the linear time trend, that without confounders was -0.003 (already statistically insignificant), was reduced even more, what tells us that the small part of the linear time trend captured in the first model was the effect of the evolution of these added confounders.

That said, the results still make me believe that the economic activity, proxied by GDP per capita, has a causal effect in the CO2 emissions. One possible mechanism (and so a bad confounder that one should not add in the regressions) can be the number of industrial firms. Industrial firms are highly pollutant, but they are also the main source of economic growth and production. Another possible mechanism is the individual behaviour that comes with an increase in GDP. Economic growth brings purchase power, and purchase power is often (mainly in the period considered) associated with pollutant behaviours (cars, deodorants, etc). It is true that in modern days people are getting everyday more aware of the impact of their behaviour, mainly in countries with high GDP, but that is a process that is far from being over.

Conclusion

Before making my final remarks, I would like to talk about some robustness tests I have done, mainly through weighted and unweighted regressions. All the regressions showed above were weighted on population size. Countries in the panel data have radically different population sizes, and in my understanding, accounting for such differences is significant when studying the relationship between CO2 emissions and GDP per capita. The relationship between CO2 emissions and GDP per capita in a country like China or in a country like Luxembourg have very different impacts that one should consider. That would not be a problem if both regressions would give us the same results, but the coefficients are almost doubled in the weighted regressions. For the reasons former mentioned, I consider the weighted regressions a better way to modelling the relationship of interest.

Concluding, it is obvious that there is a strong link between CO2 emissions and GDP per capita, since every model with every kind of specification has shown a statistically significant and a considerable sized coefficient. Whether that means a causal relationship or not, we need to be more careful. As shown, the introduction of four confounders have drastically dropped the size of the coefficient (even though it did not affect its significance). One can think about more confounders, and that might reduce even more the GDP coefficient. In theory, it is even possible that such coefficient is negative. However, after all the regressions (mainly the FD and FE models that account already for a lot of unobserved confounders), and given the mechanisms that I previously stated, my confidence is that economic activity, in the terms that we had between 1995 and 2018, does cause an increase in CO2 emissions.

Appendix

1) OLS with Confounders

	(1)	(2)	
	Year = 1995	Year = 2018	Standard errors in parentheses
VARIABLES	Natural logarithm of CO2 Emissions per Capita	Natural logarithm of CO2 Emissions per Capita	** p<0.01, * p<0.05
ln_gdp_per_capita	0.675**	0.797**	
	(0.062)	(0.047)	
ln_population	0.655**	0.916**	
	(0.136)	(0.102)	
ln_urban_population	0.315*	0.037	
	(0.133)	(0.097)	
ln_forest_area	0.076**	0.050**	
	(0.024)	(0.018)	
renewable_energy	-0.024**	-0.020**	
	(0.002)	(0.002)	
Constant	-11.605**	-12.744**	
	(0.750)	(0.587)	
Observations	167	167	
R-squared	0.957	0.970	

2) FD with no lags

	(1)	
	FD With No Lags	Robust standard errors in parentheses
VARIABLES	FD Natural logarithm of CO2 Emissions per Capita	** p<0.01, * p<0.05
d_ln_gdp_per_capita	0.355**	
	(0.063)	
d_ln_population	0.851**	
	(0.203)	
d_ln_urban_population	0.231	
	(0.170)	
d_ln_forest_area	-0.010	
	(0.161)	
d_renewable_energy	-0.027**	
	(0.003)	
year	-0.000	
	(0.000)	
Constant	0.207	
	(0.471)	
Observations	3,841	
R-squared	0.491	

3) Fixed Effects

	(1)	
	Fixed Effects	Robust standard errors in parentheses
VARIABLES	Natural logarithm of CO2 Emissions per Capita	** p<0.01, * p<0.05
ln_gdp_per_capita	0.265**	
	(0.076)	
ln_population	0.617*	
	(0.257)	
ln_urban_population	0.523**	
	(0.194)	
ln_forest_area	0.159	
	(0.276)	
renewable_energy	-0.024**	
	(0.004)	
Constant	-11.959*	
	(6.057)	
Observations	4,008	
Number of countryid	167	
R-squared	0.926	

4) Long Difference Model

	(1)	
	Long Difference Model	Robust standard errors in parentheses
VARIABLES	LD Natural logarithm of CO2 Emissions per Capita	** p<0.01, * p<0.05
d_ln_gdp_per_capita	0.350**	
	(0.067)	
d_ln_population	0.848**	
	(0.210)	
d_ln_urban_population	0.232	
	(0.176)	
d_ln_forest_area	-0.001	
	(0.168)	
d_renewable_energy	-0.027**	
	(0.002)	
Constant	-0.007*	
	(0.003)	
Observations	3,841	
R-squared	0.489	